

2014-5-9 周报

顾天瑜

1.目前形成的几个思路：

1)基于移动基站数据，分析农民工流动性

从几个依据出发，筛查可能是农民工的号码：

A. 在过年前走西面的火车线路离开，过年后又走这一线路返回的

B. 平时工作时间体现为三班倒（日-夜-休大三班，或日-中-夜-休小三班）的

C. 主要工作生活区域体现为城中村、工业区、村庄之一的

专家意见：只有一年的数据，无法体现时序上的变化，不便做产业经济学分析

暂时搁置，先初步摸索，看是否有其他比较有趣的模式

2)基于移动基站数据，分析工作时长和周末行动模式的关系

从几个依据出发，筛查上长白班的人的信息：

A. 周一至周五 09:00~11:00,14:00~17:00，一直在被标示为办公楼的基站区域内

B. 在工作日 22:00~06:00，生活区域和前述区域不一致，且标示为居民区

C. 周六或周日至少一天不符合这一生活模式的

先从春节长假过后的数据中，选一周进行检验，筛选出初步数据集。

然后，随机取这周以外的一天工作日，避开春节长假和长假前一周，比对是否符合长日班的工作模式。如果符合，则认为确实是长日班工作者，如果不符合，再

随机选一天，如果仍不符合就排除，符合则认为是长日班工作者。

最后，对他们在工作日的工作时长和在周末的行动轨迹进行可视化。寻找规律。

3)基于出租车数据，进行复杂网络分析

上一次投稿的出租车数据可视化，在交互上、分析上层次还比较低，和背后数据

结构上的重大创新不相匹配。因此，考虑将数据抽象为复杂网络进行研究和可视化。

出租车数据适合进行复杂网络化抽象有几个原因：

- A. 出租车数据经过清洗和修正，非常精确，便于抽象化。
- B. 出租车沿着道路轨迹开，不会乱跑，可以定位到抽象网络上。
- C. 决定城市交通情况的路口并不多，如果将他们设置为网络节点，一方面大大降低计算负担，一方面也能忠实地体现城市交通情况的特性（某些次要的小路与大路、小路与小路的路口受偶发情况影响比较大，不能体现网络总体性质）。

目前探索的几个思路：

- A. 基于改进 EGM 方法，探索用交通数据知道城市规划和交通政策的可能性
- B. 基于混沌同步理论，探索在实时交通情况系统中，对一个输入的偶发事件（比如一起重大交通事故），判断其对整个网络可能产生的影响，并基于此用可视化方法，辅助采取交通管理部门对应决策的可能性
- C. 基于复杂网络上的社团分割理论，从交通数据出发，划分整个城市中的生活圈。
- D. 基于博弈论和复杂网络的抗毁性分析，探索在重大国家安全事件发生时，利用实时更新的交通情况系统加上可视化分析，辅助进行对应策略的制定的可能性。（比如进行交通管制或在重要节点部署警力等）